

FOS-51305



日本国特許庁
PATENT OFFICE
JAPANESE GOVERNMENT

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
in this Office.

出願年月日
Date of Application:

1999年12月 8日

願 番 号
Application Number:

平成11年特許願第348422号

願 人
Applicant(s):

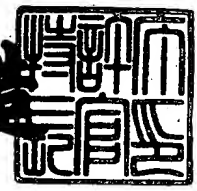
日本電気株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2000年10月27日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2000-3088634

【書類名】 特許願

【整理番号】 47201421

【あて先】 特許庁長官殿

【国際特許分類】 G06F 15/40
H04L 12/56

【発明者】

【住所又は居所】 東京都港区芝五丁目 7 番 1 号 日本電気株式会社内

【氏名】 亀谷 潤

【特許出願人】

【識別番号】 000004237

【氏名又は名称】 日本電気株式会社

【代理人】

【識別番号】 100097157

【弁理士】

【氏名又は名称】 桂木 雄二

【手数料の表示】

【予納台帳番号】 024431

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 データ検索システム及びパケット処理装置並びに制御方法

【特許請求の範囲】

【請求項 1】 所望データを記憶したデータベースと、このデータベースから選択・複写されたデータを記憶する検索用テーブルと、検索用テーブル上の各データのエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述したアドレスポインタテーブルとを備えたことを特徴とするデータ検索システム。

【請求項 2】 前記アドレスポインタテーブルは、記録媒体上に固定サイズのメモリブロックが連続して取られた領域からなり、各メモリブロック中には、前記検索用テーブル上の特定データの位置を指し示すエントリのアドレスが格納される第 1 のエリアと、この特定データに関連する次のデータのエントリのアドレスを格納したメモリブロックの位置を示すアドレスまたは当該メモリブロック自身がリストの最終であることを意味するフラグのどちらかが格納される第 2 のエリアとの 2 種類の部分が割当てられていることを特徴とする請求項 1 に記載のデータ検索システム。

【請求項 3】 所望データを記憶したデータベースと、このデータベースから選択・複写されたデータを記憶する検索用テーブルと、前記検索用テーブル上のデータのエントリアドレスの位置情報および／またはエントリアドレス相互の依存関係を記述するアドレスポインタテーブルとを備え、前記アドレスポインタテーブルは、記録媒体上に固定サイズのメモリブロックが連続して取られた領域からなり、各メモリブロック中には、前記検索用テーブル上の特定データの位置を指し示すエントリのアドレスが格納される第 1 のエリアと、この特定データに関連する次のデータのエントリのアドレスを格納したメモリブロックの位置を示すアドレスまたは当該メモリブロック自身がリストの最終であることを意味するフラグのどちらかが格納される第 2 のエリアとの 2 種類の部分が割当てられており、

検索時には、与えられた被検索データが前記検索用テーブル中に登録されているデータと一致すれば当該データを検索結果として出力し、被検索データが前記

検索用テーブル中に無かった場合には前記データベース中から一致するデータを抽出して検索結果として出力すると共に当該データを前記検索用テーブルに登録し、時間間隔等の所定条件に対応して前記検索用テーブル中の被検索頻度が低いデータを削除するとともに当該削除データに基づき前記アドレスポインタテーブルに関連づけられた全てのデータの占有記憶領域を新たに記憶が可能な領域として解放することを特徴とするデータ検索システム。

【請求項 4】 マイクロプロセッサと、ルーティングテーブルと、パケット処理を高速化するための検索用テーブルを記憶するメモリとを具備してなるパケット処理装置において、

前記検索用テーブル上のエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述したアドレスポインタテーブルを備えたことを特徴とするパケット処理装置。

【請求項 5】 マイクロプロセッサと、ルーティングテーブルと、パケット処理を高速化するための検索用テーブルを記憶する検索メモリが接続された検索回路とを具備してなるパケット処理装置において、

前記検索回路が使用する検索用テーブル上のエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述したアドレスポインタテーブルを備えたことを特徴とするパケット処理装置。

【請求項 6】 前記アドレスポインタテーブルを前記マイクロプロセッサ用のメインメモリ上に設けたことを特徴とする請求項 5 に記載のパケット処理装置。

【請求項 7】 前記アドレスポインタテーブルを前記検索メモリ上に設けたことを特徴とする請求項 5 に記載のパケット処理装置。

【請求項 8】 前記アドレスポインタテーブルは、固定サイズのメモリブロックが連続して取られた領域からなり、各メモリブロック中には、それが指し示す前記検索メモリ上のエントリアドレスを格納するエリアと、関連する次のエントリのアドレスを格納したメモリブロックのアドレスまたは自身がリストの最終を意味するフラグを格納するエリアとの 2 種類の部分からなることを特徴とする請求項 5 ～ 7 に記載のパケット処理装置。

【請求項 9】 請求項 5～8 のいずれか 1 項に記載の packets 処理装置の制御方法であって、ルーティングテーブルの内容が変更された場合にマイクロプロセッサが、変更となるエントリ中の登録内容に基づきアドレスポインタテーブルを参照し、該当するエントリ及びこのエントリに対応づけられたエントリが検索メモリ中のどこに登録されたかを調べ、これらのエントリの削除を行う過程を含む packets 処理装置の制御方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、電子データの検索システム及び packets 通信ネットワークにおいて packets の伝送を行うための packets 交換装置等の検索システムを含み構成されている packets 処理装置とその制御方法に関する。

【0002】

【従来の技術】 電子機器や通信システム等では、様々な形態でデータ検索システムが用いられている。これらのデータ検索システムは、一般にマイクロプロセッサと、被検索データ（データベース）を記憶した記憶媒体を含み構成される。記憶媒体としては各種半導体メモリやその他の電氣的記憶素子あるいはフロッピーディスク、ハードディスク、光磁気ディスク等の既知の外部記憶媒体が用いられている。

【0003】

こうしたデータ検索システムにおいて、検索動作を高速化するために常にデータベース全体を検索するのを回避するために、本来のデータベース以外に少容量の検索用テーブル（副データテーブル）を設けておき（必要に応じてはより高速動作可能な媒体に設ける）、検索に際しては先ずこの検索用テーブルを検索し該当データが無い場合にのみデータベースを検索するようにして検索時間の短縮を図る手法が知られている。このようなシステムの場合、システムの稼働中に一度データベースから検索されたデータを検索用テーブルにも記憶（登録）していくようにする。或いは前もって使用頻度が高いデータ群が予想できるようなケースであればこれらを予め検索用テーブルに記憶させておく。

【0004】

上記検索用テーブルに記憶させるデータ量は、あまり多くの量を設定すると検索の高速化に繋がらないので、使用頻度の高い少量のデータのみを検索用テーブルに記憶させておくように配慮が重要となる。

【0005】

また、検索用テーブル用にはデータベース用よりも高速動作の半導体メモリを使用することがあるが、この場合には一般に記憶素子の価格もより高価でコストの面からも検索用テーブルの記憶容量は少ないことが望ましいことになる。こうした理由から、検索用テーブル上のデータは使われないうまま時間が経過したもの使用頻度の低いものから削除されてゆく。

【0006】

上述したようなデータ検索システムを用いている例としてパケット交換装置を挙げてその動作を詳述する。一般にインターネットに代表されるパケット通信ネットワークでは、パケットの伝送はルータ等のパケット交換装置によって、ヘッダ部分のアドレス情報に基づき、パケット単位で適切なネットワークへ転送する事により行われている（ルーティング）。このパケット単位でのルーティング処理という性質上、ルータのパケット処理はソフトウェアで実現される事が多かった。

【0007】

しかしながら、最近のルータには、従来のパケットルーティング処理やフォワーディング処理以外に、IPsec への対応といった付加的な処理についての高速化も同時に求められている。これらの処理においては、基本的には個別のIPパケットをIP Destination AddressやIP Source Address 等に基づき分類（Micro Flow）し、受信パケットの外側にカプセル化ヘッダを新たに付加する事により、アグリゲートしたIPトラフィックの単位（Aggregated Flow）を作り出し、その新しいIPトラフィックに対する統一的な処理を実行する事が求められている。この要求に対応させたパケット交換装置が、本出願人による特願平11-098140号（以下、既提案と記す）の明細書及び図面にも記載されている。

【0008】

この既提案（図7参照）は、検索ハードウェア（サーチ処理部および対応した検索用メモリ）を用いて、比較的単純な構成で上述したようなIPsec 処理等の高速化を実現したものである。すなわち、パケット処理高速化の要求を満たすために前述したような検索用テーブルとしてのIPフローテーブル及び専用の検索回路（サーチ処理部）を備えたパケット処理装置を提案している。

【0009】

なお、この既提案はセキュリティ処理部を有しておりIPsecに準じたパケットデータの暗号化／復号化処理を実装する際の処理高速化に繋げるため、マイクロプロセッサに依らずセキュリティ処理手段により処理する構成をも提案しているがこの点については以下の説明中では本発明の明瞭化の観点から触れない。

【0010】

上記既提案に記載されている装置に準じた、検索用テーブル（IPフローテーブル17a）及び検索回路（サーチ処理部16）を備えたパケット交換装置（レイヤ3スイッチ）の構成要部を示すブロック図を図7に示す。この装置の検索動作は、サーチ処理部（検索回路）を用いて検索メモリ上の検索用テーブルであるIPフローテーブルを検索する事によって行われ、マイクロプロセッサ上のソフトウェアによる検索と比べ高速化が図られている。

【0011】

図7のパケット交換装置は、マイクロプロセッサ11と、ルーティングテーブル12aがおかれるメインメモリ12と、プロセッサバス13と、マイクロプロセッサ11に代わってパケット処理を実行するパケット処理部14、パケットメモリ15、サーチ処理部16及び検索メモリ17中のIPフローテーブル17aと、下位レイヤ処理部20と、パケット処理部14と下位レイヤ処理部20とを接続するスイッチファブリック18とを備えている。なお、セキュリティ処理部19がスイッチファブリック18を介してパケット処理部14と下位レイヤ処理部20に接続されている。

【0012】

マイクロプロセッサ11は装置全体の制御等を行うと共に、パケット処理をソフトウェアで実行する。メインメモリ12はマイクロプロセッサ11用のプログ

ラムや、ルーティングテーブル等を格納しておくためのメモリである。

【0013】

ルーティングテーブル12aはメインメモリ12上に置かれ、マイクロプロセッサ11のソフトウェア処理に基づき作成される、受信パケットのIP Destination Address中のNetwork Prefixを検索キーとし、出力ポート番号やネクストホップのレイヤ2アドレスを内容とするデータベースである。

【0014】

また、サーチ処理部16はマイクロプロセッサ11もしくは、パケット処理部（パケット転送用コントローラ）によって起動され、IPフローテーブル17a（検索用テーブル）に対して特定の検索キーに基づく検索を実行し、その結果を起動元に通知する。検索メモリ17に格納されたIPフローテーブル17aはマイクロプロセッサ11のソフトウェアにより登録エントリの管理が行われ、そのエントリに対するサーチ処理部16からの検索が実行される検索用テーブルである。プロセッサバス13が上述した各構成要素を接続している。

【0015】

図8は、上述装置におけるパケット処理（セキュリティ関連処理については省略）の主要な流れを示すフローチャートであり、また図9はマイクロプロセッサ11によるルーティング処理の流れを示すフローチャートである。両フローチャートはどちらも前述既提案に開示されているものである。フローチャートの説明、上述装置の詳細説明や説明した部分以外の構成、また装置動作の詳細等については既提案に説明されているのでここでの説明は省略する。なお、本発明に関連深い部分やその動作については後述する本発明の実施例中の同等部分に関連づけて説明する。

【0016】

上記既提案の開示技術によれば、データ検索システムを含み構成されたパケット交換装置において、カプセル化する前のオリジナルパケット（Micro Flow）とカプセル化後のパケット（Aggregated Flow）の転送処理を、同一ハードウェアにより実行する事により装置構成の単純化ならびに高速化が実現可能となっている。

【0017】

【発明が解決しようとする課題】

ところで、上述開示技術においてはアグリゲートしたIPトラフィック（カプセル化後のIPパケット）に対する処理内容が変更となった場合（セキュリティアソシエーションの失効等）の、検索ハードウェアに対する変更通知に関連して課題を残している。

【0018】

すなわち、検索メモリ17上の検索用テーブル（IPフローテーブル17a）の各エントリはIPトラフィックの出現の順に作成されるため、マイクロプロセッサ11が、アグリゲートしたIPトラフィックへの変更を元の各エントリに反映させるためには検索メモリ17に対し該当するエントリをテーブル検索しながら発見しなければならない。

【0019】

上記過程は、アグリゲートするIPトラフィックの数が数個単位であれば問題にならないが、数十、数百のエントリを発見するために各エントリを舐潰しに検索する事になると、このための検索時間が無視できなくなる。更に、その検索中は検索回路側からの検索動作が待たされる事となるため、折角のハードウェア検索導入による処理高速化の妨げとなり、装置全体としてのスループットの低下を引き起こす原因となってしまうとの問題点を有している。

【0020】

ちなみに、個々のIPフロー（Micro Flow）の検索メモリ17におけるエントリアドレスを、直接ルーティングテーブル12a等のデータベースに記録するとの解決手法も考えられるが、好ましい方法とはいえない。Aggregated Flow に対しどの個数のMicro Flowが属するかは一般的に不定であり、このデータベース用のサイズの見積りが困難となったり、確保したデータベース領域に無駄が生じるという問題が生じてしまう。

【0021】

上記説明では、パケット処理装置としてのルータを例に挙げているが、この装置に限らず、検索専用の検索用テーブルを設けておき、一度検索対象となったデ

ータについてはこのデータを検索用テーブル側にも記憶しておき、新たな検索に際しては先ず当該被検索データが検索用テーブルに在るか否かを検索し、無い場合にのみ基となるデータベースを検索するようにした検索システムは各種用途で用いられている。

【0022】

そして、このような検索システムにては高速化効果を維持する目的からも、また記憶媒体のコストという面からも検索用テーブルに保持させるデータ量は制限され、当然ながら使用頻度や記憶されてからの経過時間あるいは登録順序に応じて使用頻度の低いデータについては順次削除していかなければならない。複数データが参照関係にある複合構造においては、データ削除に伴い検索用テーブルの整合性を保つための処理に時間がかかり高速動作の妨げになるという問題点は共通して生じる。

【0023】

本発明は、検索システムにおける上述したような事情に鑑みて創案されたものでありその目的は、前述したような検索システムにおいて高速検索のための検索用テーブルからデータ削除を行った場合でも簡単・短時間の処理で当該テーブルの整合性を保つことによって、検索システムを含み構成されたシステム全体の処理速度が低下することを防止できる新規な検索システムを提案することを目的とする。

【0024】

また、このような検索システムの好適な用途装置の一つである、マイクロプロセッサと、パケット処理を高速化するための検索用テーブルを記憶する検索メモリが接続された検索回路とを具備してなるパケット処理装置において、パケットのルーティング処理の一環として用いられる検索用テーブルメモリの管理方法を簡略化して、テーブル管理処理時間を大幅に短縮し、もってパケット処理性能の低化を防止できるパケット処理装置とその制御方法を提案することを目的としている。

【0025】

【課題を解決するための手段】 上記課題を解決するために本発明では、デ

ータ検索システムを、所望データを記憶したデータベースと、このデータベースから選択・複写されたデータを記憶する検索用テーブルと、検索用テーブル上の各データのエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述したアドレスポインタテーブルとを備えた構成とする。

【0026】

更には、前記アドレスポインタテーブルが、記録媒体上に固定サイズのメモリブロックが連続して取られた領域からなり、各メモリブロック中には、前記検索用テーブル上の特定データの位置を指し示すエントリのアドレスが格納される第1のエリアと、この特定データに関連する次のデータのエントリのアドレスを格納したメモリブロックの位置を示すアドレスまたは当該メモリブロック自身がリストの最終であることを意味するフラグのどちらかが格納される第2のエリアとの2種類の部分が割当てられている構成とする。

【0027】

また、データ検索システムを、所望データを記憶したデータベースと、このデータベースから選択・複写されたデータを記憶する検索用テーブルと、前記検索用テーブル上のデータのエントリアドレスの位置情報および／またはエントリアドレス相互の依存関係を記述するアドレスポインタテーブルとを備え、前記アドレスポインタテーブルは、記録媒体上に固定サイズのメモリブロックが連続して取られた領域からなり、各メモリブロック中には、前記検索用テーブル上の特定データの位置を指し示すエントリのアドレスが格納される第1のエリアと、この特定データに関連する次のデータのエントリのアドレスを格納したメモリブロックの位置を示すアドレスまたは当該メモリブロック自身がリストの最終であることを意味するフラグのどちらかが格納される第2のエリアとの2種類の部分が割当てられており、検索時には、与えられた被検索データが前記検索用テーブル中に登録されているデータと一致すれば当該データを検索結果として出力し、被検索データが前記検索用テーブル中に無かった場合には前記データベース中から一致するデータを抽出して検索結果として出力すると共に当該データを前記検索用テーブルに登録し、時間間隔等の所定条件に対応して前記検索用テーブル中の被検索頻度が低いデータを削除するとともに当該削除データに基づき前記アドレス

ポインタテーブルに関連づけられた全てのデータの占有記憶領域を新たに記憶が可能な領域として解放するようにする。

【0028】

また、本発明のパケット処理装置では、マイクロプロセッサと、ルーティングテーブルと、パケット処理を高速化するための検索用テーブルを記憶するメモリとを具備してなるパケット処理装置において、検索用テーブル上のエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述したアドレスポインタテーブルを備える。検索用テーブルを記憶する検索メモリが接続された検索回路を具備するようにしても良い。また、前記アドレスポインタテーブルを前記マイクロプロセッサ用のメインメモリ上に、あるいは前記検索メモリ上に設ける。

【0029】

また、前記アドレスポインタテーブルを、固定サイズのメモリブロックが連続して取られた領域からなり、各メモリブロック中には、それが指し示す前記検索メモリ上のエントリのアドレス格納エリアと、関連する次のエントリのアドレスを格納したメモリブロックのアドレスまたは自身がリストの最終を意味するフラグを格納するエリアとの2種類の部分からなるようにする。

【0030】

そして本発明方法では、上述のごときパケット処理装置において、ルーティングテーブルの内容が変更された場合にマイクロプロセッサが、変更となるエントリ中の登録内容からアドレスポインタテーブルを参照し、該当するエントリ及びこのエントリに対応づけられたエントリが検索メモリ中のどこに登録されたかを調べ、これらのエントリの削除を行う過程を含むように制御する。

【0031】

【発明の実施の形態】 本発明によれば、データ検索システムを、データベースと、検索用テーブルと、検索用テーブル上の各データのエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述したアドレスポインタテーブルとを備えた構成とする。

【0032】

更には、前記アドレスポインタテーブルは、記録媒体上に固定サイズのメモリブロックが連続して取られた領域からなり、各メモリブロック中には、前記検索用テーブル上の特定データの位置を指し示すエントリのアドレスが格納される第1のエリアと、この特定データに関連する次のデータのエントリのアドレスを格納したメモリブロックの位置を示すアドレスまたは当該メモリブロック自身がリストの最終であることを意味するフラグのどちらかが格納される第2のエリアとの2種類の部分が割当てられた構成とすることができる。

【0033】

或いは、データ検索システムを、データベースと、検索用データテーブルとアドレスポインタテーブルとを備え、アドレスポインタテーブルは、記録媒体上に固定サイズのメモリブロックが連続して取られた領域からなり、各メモリブロック中には、前記検索用テーブル上の特定データの位置を指し示すエントリのアドレスが格納される第1のエリアと、この特定データに関連する次のデータのエントリのアドレスを格納したメモリブロックの位置を示すアドレスまたは当該メモリブロック自身がリストの最終であることを意味するフラグのどちらかが格納される第2のエリアとの2種類の部分が割当てられていて、アドレスポインタテーブルに検索用テーブル上のデータのエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述しておき、検索時には、与えられた被検索データが前記検索用テーブル中に登録されているデータと一致すれば当該データを検索結果として出力し、被検索データが前記検索用テーブル中に無かった場合には前記データテーブル中から一致するデータを抽出して検索結果として出力すると共に当該データを前記検索用テーブルに登録し、一定時間間隔等の所定条件に対応して前記検索用テーブル中の被検索頻度が低いデータを削除するとともに当該削除データに基づき前記アドレスポインタテーブルにより指示され参照される全てのデータの占有記憶領域を新たな記憶が可能な領域として解放するように構成する。

【0034】

上述のように、本発明では構成要素または属性が異なる検索キーの検索対象エントリが複数混在しており、かつ一群のエントリ間それぞれに依存関係がある場

合に、アドレスポインタテーブルをメモリ上に別途設けるようにしている。これにより、相互に依存関係にある検索対象エントリに対する操作をアドレスポインタテーブルを利用することによって高速かつ簡便なものにすることが可能となっている。

【 0 0 3 5 】

上記検索動作を全て単一の制御部（マイクロプロセッサ等）により行っても良いが、検索用テーブルを検索するための専用のハードウェアを別途備えた構成とすることもできる。検索用テーブル及びアドレスポインタテーブルを記憶する記

憶媒体には、半導体メモリの他ハードディスク等の記憶装置を用いることができる。検索用テーブルには、より高速な記憶媒体を用いるようにしても良く、データ検索システムの動作を更に高速化することができる。

【 0 0 3 6 】

また本発明では、構成要素または属性が異なる検索キーの検索対象エントリが複数混在し、かつそれぞれのエントリ間に依存関係があるようなデータ検索システムを含み構成されているルータその他のパケット処理装置において、検索用テーブル上のエントリアドレスの位置情報および／またはエントリアドレス相互の依存関係を記述したアドレスポインタテーブルをメモリ上に別途設けるようにする。これにより、相互に依存関係を持つ検索対象エントリの登録・削除といった操作をアドレスポインタテーブルを利用することによって高速かつ簡便なものにすることができ、結果、パケット操作も高速に行える。

【 0 0 3 7 】

上記検索動作は全て単一のマイクロプロセッサが行うように構成しても良いが、検索用テーブル上の検索のために専用のハードウェアを備えた構成としてより高速化を図ることもできる。

【 0 0 3 8 】

パケット処理装置は高速動作を要求されるため本発明に係るデータ検索システムの好適な用途の一つであるが、これに限らずデータ検索システムにおいて本発明を適用するならば、一般に検索用テーブル上でのデータ登録・削除の繰り返し

動作を含めた高速な検索動作とその維持が保証される。

【0039】

【実施例】

以下、パケット交換装置を実施例とし図面を参照して本発明について詳細に説明する。本発明の一実施例を示している図1は、テーブル検索を必要とし、本発明に係るデータ検索システムを含み構成された本発明に係るパケット処理装置としてのルータ装置（パケット交換装置）の構成要部を示す概略ブロックダイアグラムである。この実施例は、データベースとしてのルーティングテーブル12aと、検索用テーブルとしてのIPフローテーブル17aそして、アドレスポインタテーブルAPTを備えたデータ検索システムの実施例であると同時にパケット処理装置の実施例でもある。図示装置の大部分の構成は先に図7に示した装置と共通している。

【0040】

すなわち、図1において、11はマイクロプロセッサ、12はメインメモリ、12aはルーティングテーブル、13はプロセッサバス、14はパケット処理部、15はパケットメモリ、16はサーチ処理部、17は検索メモリ、17aはIPフローテーブル、18はスイッチファブリック、19はセキュリティ処理部、そして20は下位レイヤ処理部であり、これらの構成は先の図7の装置と略同等である。

【0041】

上記各部に加えて、実施例装置はメインメモリ12にアドレスポインタテーブル；APTが設けられており、この点が従来の装置と異なっている。このアドレスポインタテーブル；APTは、ルーティングテーブル12aおよびIPフローテーブル17aとともに本発明のデータ検索システムの主要部を構成する。このデータ検索システムに依ってルータ装置（パケット交換装置）の性能向上を実現している。以下既知部分も含めて各構成について説明する。

【0042】

マイクロプロセッサ11は、装置全体の制御を行うと共に、特定のルーティングアルゴリズムとプロトコルに基づきパケットのネクストホップを決定し、その

処理結果をルーティングテーブルに反映する等の、パケット処理をソフトウェアで実行する。メインメモリ 12 はマイクロプロセッサ 11 自身のソフトウェアや、ルーティング処理のためのデータベースやテーブルを、格納しておくためのメモリである。

【0043】

ルーティングテーブル 12a はメインメモリ 12 上に置かれ、マイクロプロセッサ 1 のソフトウェア処理に基づき作成される、受信パケットの IP Destination Address 中の Network Prefix を検索キーとした、出力ポート番号やネクストホップのレイヤ 2 アドレスを内容とするデータベースである。

【0044】

そして同じくメインメモリ 12 上に置かれた アドレスポインタテーブル ; ATP にはルーティングテーブル 12a 中のエントリと IP フローテーブル 17a (検索用テーブル) 上のエントリの依存関係、格納アドレス情報等が格納される。

【0045】

検索回路としてのパケット処理部 (転送用コントローラ) 14 は、ネットワークから受信したパケットに対して、IP ヘッダ処理や出力先ネットワークインターフェイスへの転送処理を行う。このパケット処理部 14 にはパケットメモリ 15 が接続されており、またプロセッサバス 13 を介して前述のマイクロプロセッサ 11 等と接続されている。また、パケット処理装置 14 にはスイッチファブリック 18 を介してセキュリティ処理部 19、そして図示していないネットワークインターフェイスへとつながった下位レイヤ処理部 20 がそれぞれ接続されている。

【0046】

サーチ処理部 (検索回路) 16 は前記プロセッサバス 13 に接続されており、マイクロプロセッサ 11 もしくは、パケット処理部 14 によって起動されて、IP フローテーブル 17a に対して特定の検索キーに基づく検索を実行し、その結果を起動元に通知する。

【0047】

検索メモリ 1 7 はサーチ処理部 1 6 に接続されていて IP フローテーブル 1 7 a を格納する。この検索メモリ 1 7 上に置かれている IP フローテーブル 1 7 a は、IP Source Address と IP Destination Address とをサーチキーとして、マイクロプロセッサ 1 1 によるルーティング処理の結果である IP フローテーブルを格納したテーブルで、マイクロプロセッサ 1 1 のソフトウェアにより登録エントリの管理が行われ、そのエントリに対してサーチ処理部 1 6 からの検索が実行される。

【 0 0 4 8 】

本発明に係るデータ検索システム（メモリ管理方法）は、実施例のように検索ハードウェアを用いる事により、マイクロプロセッサ上のソフトウェアによるルーティングテーブル等のデータベースの検索処理に比べ、性能向上を図った例えばルータ等のパケット処理装置に適用して好適なものである。

【 0 0 4 9 】

ここで、簡略に図 1 の例示装置におけるマイクロプロセッサと検索ハードウェアの処理の相互関連について説明する。マイクロプロセッサ 1 1 は、その上で動作するソフトウェア処理により、特定のプロトコルおよび適宜のルーティングアルゴリズムに基づき、自らが接続されたネットワークトポロジーの学習やメトリック計算を実行し、IP Destination Address の Network Prefix を検索キーとし、出力ポート番号やネクストホップのレイヤ 2 アドレスを内容とするルーティングテーブル（主データベース） 1 2 a を、メインメモリ 1 2 上に作成する。このルーティングテーブルの更新は、特定ルーティングプロトコルに基づき随時あるいは定期的に更新される。

【 0 0 5 0 】

一方、本発明に関連するパケット処理に関し説明すると、図示されていないネットワークインタフェース回路に到着したパケットは、レイヤ 2 以下の処理が行われ、パケットメモリ 1 5 に一旦格納される。格納されたパケットの IP ヘッダ部分は、マイクロプロセッサ 1 1 あるいはパケット処理部 1 4（パケット転送用コントローラ）により解析される。そして、正常パケットであると判断された場合には、まず当該パケットがサーチ処理部 1 6（検索回路）へと転送されて、サーチ

処理部の処理が起動される。

【0051】

サーチ処理部16では、受信したIPヘッダ部分から必要なデータ（IP Destination Address、IP Source Address 等）を抽出して、検索キーを組み立てる。サーチ処理部16はこの検索キーに基づき、検索メモリ17上の検索用テーブルとしてのIPフローテーブル17aを検索し、その結果として必要なデータ（出力ポート番号やネクストホップのレイヤ2アドレス、カプセル化ヘッダ等）を取り出し、起動元に通知する。サーチ処理部16の起動元は、受け取ったデータに基づきパケットのヘッダ更新等の処理や、必要ならばカプセル化後のIPヘッダに対する新たな検索処理の起動等を実行する。

【0052】

これら一連の処理が完了したパケットデータは、パケットメモリ上から読み出され、送信すべきネットワークワークインタフェース回路におけるレイヤ2以下の処理を経て、ネットワーク上に転送されることになる。

【0053】

なお、例示のパケット交換装置は上記ルーティング処理とともに、必要に応じてパケット単位にデータの暗号化／復号化等のセキュリティ処理を併せて実行するようになっているが、これらの処理を含め装置のパケット処理動作については先の既提案に説明があり、ここでの詳細な説明は省略する。

【0054】

図2は、IPフローテーブル17aの既提案にて例示されている一例を示す説明図であり、IPフローエントリごとに、サーチキー（IP Source Address、IP Destination Address）と、ルーティング処理の結果であるMAC発信元アドレス、MAC宛先アドレス及び出力物理ポートのポート番号とが登録されている。なお、図示例では必要に応じてセキュリティ情報も格納される。

【0055】

図3は、本発明におけるアドレスポインタテーブル；APTの構造、ならびにルーティングテーブル12a中のエントリと、IPフローテーブル17a（検索メモリ17）中のエントリの相互関係を示す説明図である。

【0056】

ルーティングテーブル 1 2 a の各個別データの末尾には 2 つのエリア (PTR1, PTR2) が確保され、自身のエントリがアグリゲートされた IP フローである場合には、PTR1 は検索メモリ 1 7 中の IP フローテーブル 1 7 a に書かれたアドレスを示し、PTR2 はそのアグリゲートされたフローに属する元のマイクロフローを示すアドレスポインタテーブル ; APT の Linked-List 構造の先頭アドレスを示すようになっている。一方、ルーティングテーブル 1 2 a のエントリ自身がマイクロフローの場合には、この 2 つの PTR1、PTR2 のエリアには NULL が書き込まれる。

【0057】

アドレスポインタテーブル ; APT は実施例ではメインメモリ 1 2 上に置かれていて、固定サイズ (例えば 4 バイト) のメモリブロックが複数個連続して取られた領域からなる (図 3 参照)。各メモリブロックは、順序立てて一巡して関連付けられて使用領域の開始ブロックと終端ブロックを管理することで空き領域が管理される。

【0058】

各メモリブロック中には、それが指し示す IP フローテーブル 1 7 a 上のエントリのアドレス格納用の第 1 のエリアと、当該メモリブロック自身がリストの最終を意味するフラグを (例 ; AD4) 或いはもし有れば関連する次のエントリのアドレスを (第 1 のエリアに) 格納しているメモリブロックのアドレスを格納する (例 ; AD1) ための第 2 のエリアとの 2 種類の部分が割り当てられている。

アドレスポインタテーブル 1 7 a 中の Linked-List 構造の管理は、アグリゲートされた IP フローに属するマイクロフローが検出された時に、マイクロプロセッサ 1 1 によって PTR1 の値に基づき IP フローテーブル上の Linked-List 構造を辿り、その構造の末尾にマイクロフローのエントリアドレスを追加する事によって行われる。

【0059】

本発明の特徴であるアドレスポインタテーブル ; APT に関連した処理について、図 3 を用いて更に詳細に説明する。検索メモリ中の IP フローテーブル上のエントリは、受信パケットのトラフィック特性に応じて受信順に順次追加される

ため、メモリ中のエントリの順序は常に不定となる。

【0060】

一方、ルータにおけるダイナミックルーティング動作の結果、ルーティングテーブル 12 a の内容が変更された場合には、対応する IP フローテーブル上のエントリをその度に変更しないと整合性が崩れ、結果として転送パケットのルートが最適化されないか、または以後のリンク上でパケットが廃棄されてしまう可能性がある。

【0061】

したがって、実施例ではダイナミックルーティングの結果等により、ルーティングテーブル（データベース）の内容を変更する必要がある場合には、マイクロプロセッサ 11 はまずルーティングテーブル中の変更となるエントリ末尾の PTR2 の情報を基に、アドレスポインタテーブルを参照し、該当するエントリが IP フローテーブル（検索メモリ）中のどこに登録されたかを調べる。

【0062】

アドレスポインタテーブルの情報は Linked-List 構造により、個々のエントリの IP フローテーブル中での存在アドレスおよび、次のエントリの有無と、次のエントリが存在する場合にはその存在アドレスの格納先を示す構造になっている。したがって、マイクロプロセッサ 11 は、このアドレスポインタテーブルを順に辿っていく事により、IP フローテーブル 17 a 中に展開された個々のエントリを検索動作無しで突き止める事が可能となる。従って、これらのエントリの削除・変更（メモリ領域の解放）を高速に行う事ができ装置の高速化につながる。

以下では本発明の実施例の特に検索に関わる動作を図 1 および図 4 ～図 6 を参照して、詳細に説明する。図 4、図 5 は実施例装置におけるアドレス検索に関連する処理の一例を示すフローチャート、図 6 はエントリ削除処理の一例を示すフローチャートである。

【0063】

実施例装置がパケットを受け取ると検索処理が行われる。すなわち、図 4 に示すように、受信した IP ヘッダ部分から IP Destination Address、IP Source Address を所定の検索キーとして抽出する (S101)。次いでこの検索キーを基に検索用

テーブル（IPフローテーブル）を検索する(S102)。

検索の結果、該当データがあれば（エントリにヒット）出力ポート番号等の情報を読み出し(S103,S104)、サーチ処理部の起動元に情報を通知する(S105)。

【0064】

過程(S103)でIPフローテーブル中に該当データが見出せない場合には、図5に示す一連の処理を行う。

【0065】

まず、IPヘッダ等の情報がマイクロプロセッサへ転送され(S106)、マイクロプロセッサによってソフトウェアによるルーティング処理が実行される(S107)。すなわち、ソフトウェア処理によりIPヘッダ部分の解析が行われ、必要な検索キーが組立てられてルーティングテーブルを検索する。得られたルーティング結果はIPフローテーブルに登録される(S108)。

【0066】

続いてルーティング結果がマイクロフローであるか否かが判定され(S109)、マイクロフローであればルーティングテーブルのPTR2を読み出し(S110)、PTR2の情報に基づきアドレスポインタテーブルをたどりリストの末尾を特定する(S111)。そして、このリスト末尾のブロックにIPフローテーブル上のデータアドレスを追加して(S112)検索処理が終了する。過程(S109)でマイクロフローではないと判定された場合には、ルーティングテーブル上の該当データのPTR1にIPフローテーブル上のアドレスを書込み(S113)、検索処理を終了する。

【0067】

上述の一連の処理の結果、次回以降は一度受信した同じIPヘッダ情報を持つ受信パケットに対しては、サーチ処理回路によるIPフローテーブルの検索にて全てヒットするため検索が高速化され、高速なルーティング従って高速なパケット転送動作が可能となる。

【0068】

次に、図6のフローチャートに示す不要となったデータの削除処理について説明する。まず、ルーティングテーブル中の該当エントリのPTR2を読みだす(S201)。そしてPTR2の情報により指示されるアドレスポインタテーブルのリストに格

納されているアドレスを読み出し(S202)、IPフローテーブル上の該当アドレスに登録されているエントリを削除する(S203)。アドレスポインタテーブルの当該メモリブロックがリストの末尾であるかを調べ(S204)、末尾ではない場合にはアドレスポインタテーブルのリストを1つ進み(S205)、次のメモリブロックの格納内容に従って過程(S202)・過程(S203)を繰り返す。

【0069】

リストの末尾に達した場合には該当リストのチェーンを解放する(S204,S206)。次いでルーティングテーブル中の該当エントリのPTR1を読み出し(S207)、IPフローテーブル上の該当アドレスに登録されているエントリを削除する(S208)。更にルーティングテーブル中の該当エントリを削除または変更する(S209)。これにより関連づけられていた一連のデータが、各テーブル中から効率的に削除される。

【0070】

以上説明したように実施例の装置では、次のような効果を得ることができる。すなわち、マイクロプロセッサが管理するルーティングテーブルのエントリと、サーチ処理回路が検索するIPフローテーブルのエントリとの関係を表わすアドレスポインタテーブルを持つ事により、カプセル化によるIPフローのアグリゲート処理等を行う際のメモリ管理方法が簡略化できる。

【0071】

これにより、マイクロプロセッサによる検索メモリ上のIPフローテーブルのサーチが不要となり、ダイナミックルーティング等を使用時のソフトウェアによるテーブル管理の処理時間が大幅に短縮できる。

【0072】

また、上記の管理処理短縮効果により、サーチ処理回路の処理待ち合せ時間も短縮できる結果、装置全体で見たパケット転送等のパケット処理のスループット低下が防止できる。

【0073】

また、アドレスポインタテーブルを利用し、検索キーの構成または属性が異なるエントリの依存関係を表現できるため、IPsec やDiffserve 等のIPフローのA

グリゲート処理が必要なパケット処理装置の構成を簡略化する事が可能となる。

上記実施例装置の効果は、本発明のデータ検索システムの高速化に起因しており、このようなデータ検索システムは汎用性がありパケット処理装置以外にも広く用いることができる。

【0074】

すなわち、本発明によればデータ検索システムを、所望データを記憶したデータベース（実施例ではルーティングテーブル）と、このデータベースから選択・複写されたデータを記憶する検索用テーブル（実施例ではIPフローテーブル）と、検索用テーブル上の各データのエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述したアドレスポインタテーブルとを備えた構成とすることにより、検索処理の高速化が達成でき各種用途に利用し得る。

【0075】

【発明の効果】 以上説明したように、本発明のデータ検索システムによれば検索処理を高速化できこれを維持できるとの効果が得られる。また、本発明のパケット処理装置・制御方法によればパケット処理に関連して高速な検索処理を行えこれを維持することができ、結果として高速なパケット処理が行えるとの効果が得られる。

【図面の簡単な説明】

【図1】 本発明に係るパケット処理装置の一実施例としてのルータ装置の構成要部を示す概略ブロックダイアグラムである。

【図2】

本発明に係るIPフローテーブルの一例を示す説明図である。

【図3】

本発明におけるアドレスポインタテーブルの構造、データベース中のエントリ、検索用テーブル中のエントリの相互関係を示す説明図である。

【図4】

実施例装置のアドレス検索に関連する処理を示すフローチャートである。

【図5】

実施例装置のアドレス検索に関連する処理を示すフローチャートである。

【図 6】

実施例装置におけるアドレス検索に関連する処理の一例を示すフローチャートである。

【図 7】

従来のパケット処理装置の構成を示す概略ブロック図である。

【図 8】

図 7 のパケット処理装置におけるパケット処理の主要部を示すフローチャートである。

【図 9】

図 7 のパケット処理装置におけるルーティング処理を示すフローチャートである。

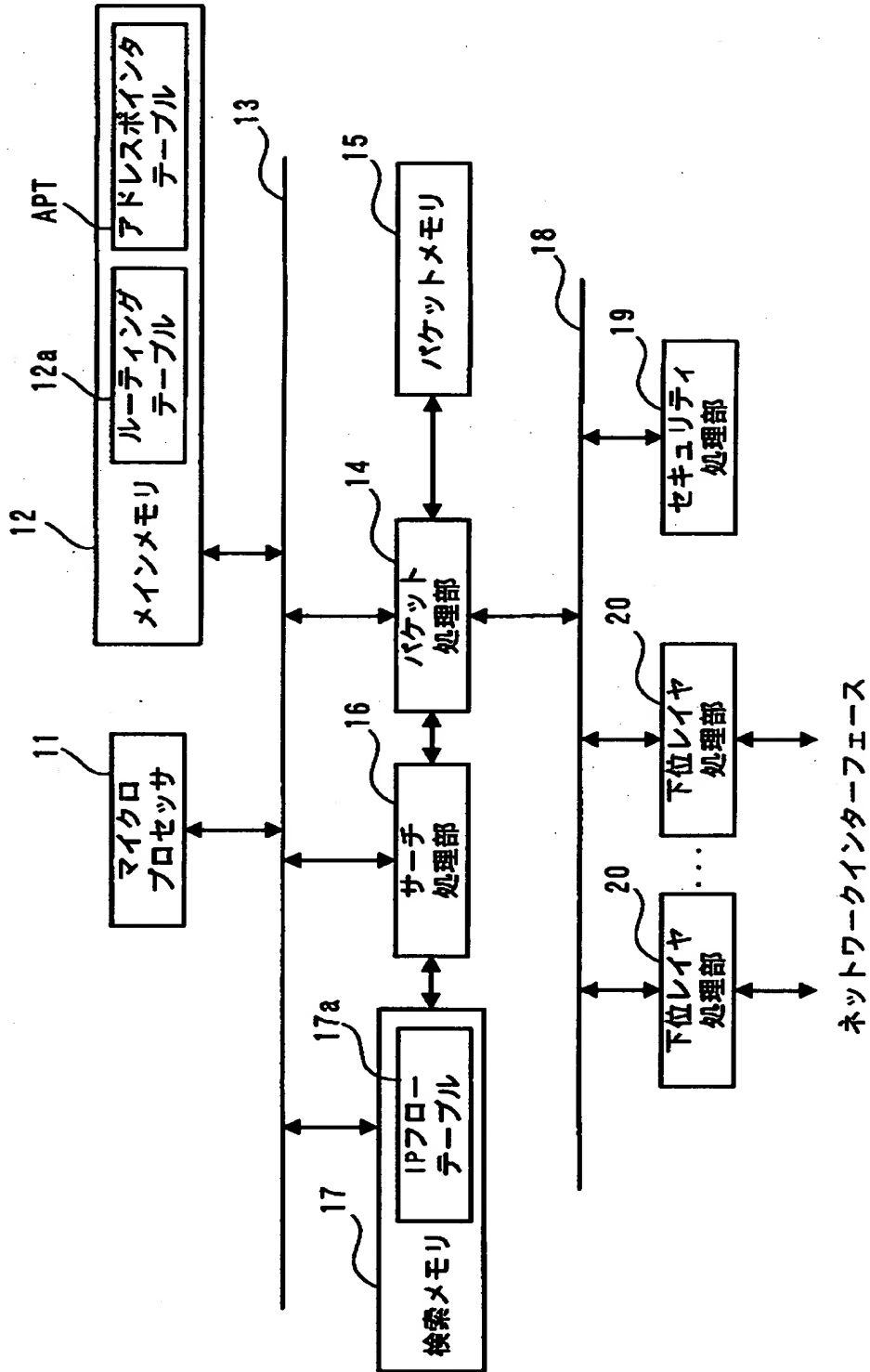
【符号の説明】

- 1 1 …マイクロプロセッサ
- 1 2 …メインメモリ
- 1 2 a …ルーティングテーブル（データベース）
- 1 3 …プロセッサバス
- 1 4 …パケット処理部
- 1 5 …パケットメモリ
- 1 6 …サーチ処理部（検索回路）
- 1 7 …検索メモリ
- 1 7 a …IPフローテーブル（検索用テーブル）
- 1 8 …スイッチファブリック
- 1 9 …セキュリティ処理部
- 2 0 …下位レイヤ処理部
- A P T …アドレスポインタテーブル

特平 1 1 - 3 4 8 4 2 2

【書類名】 図面

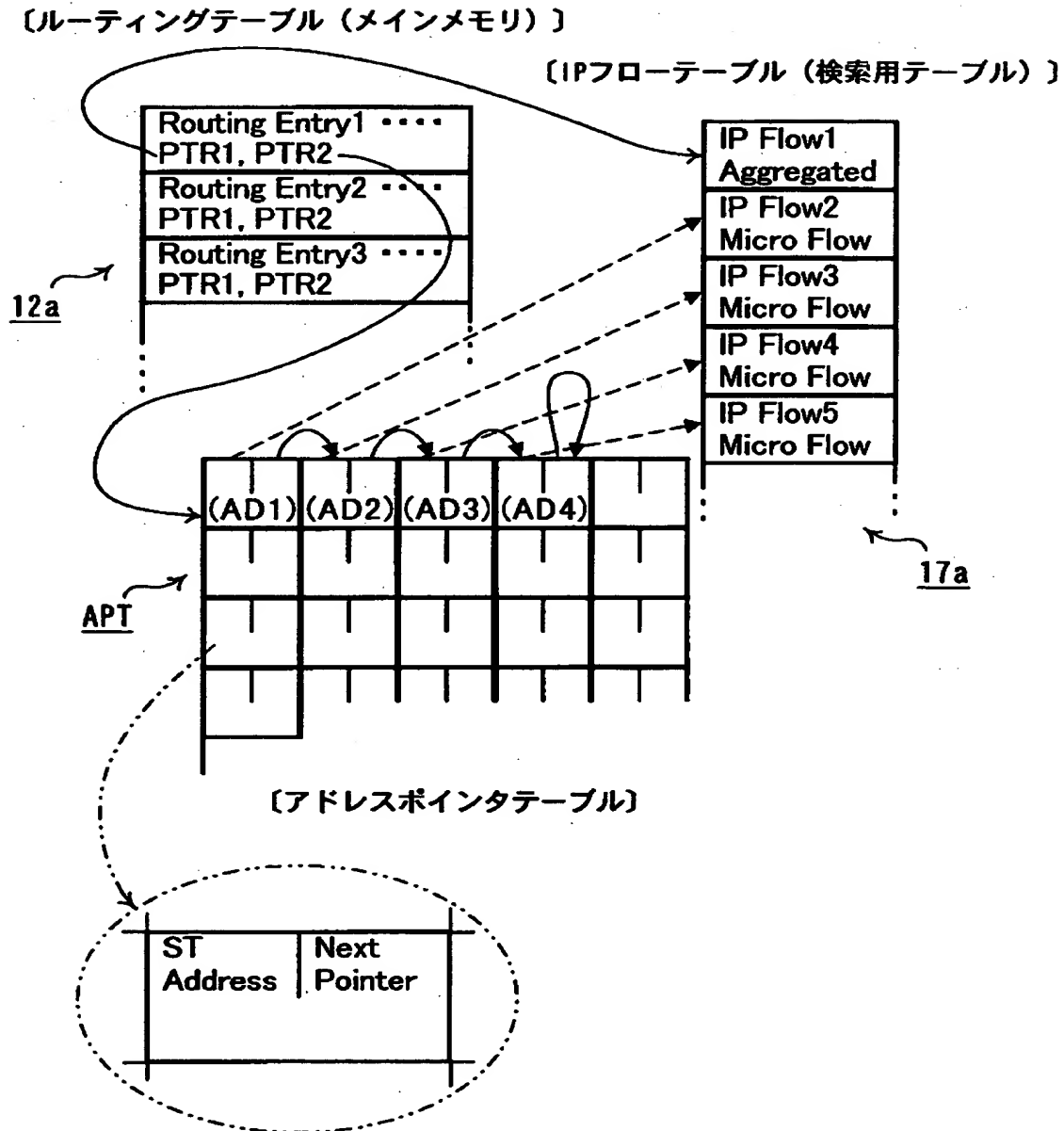
【図 1】



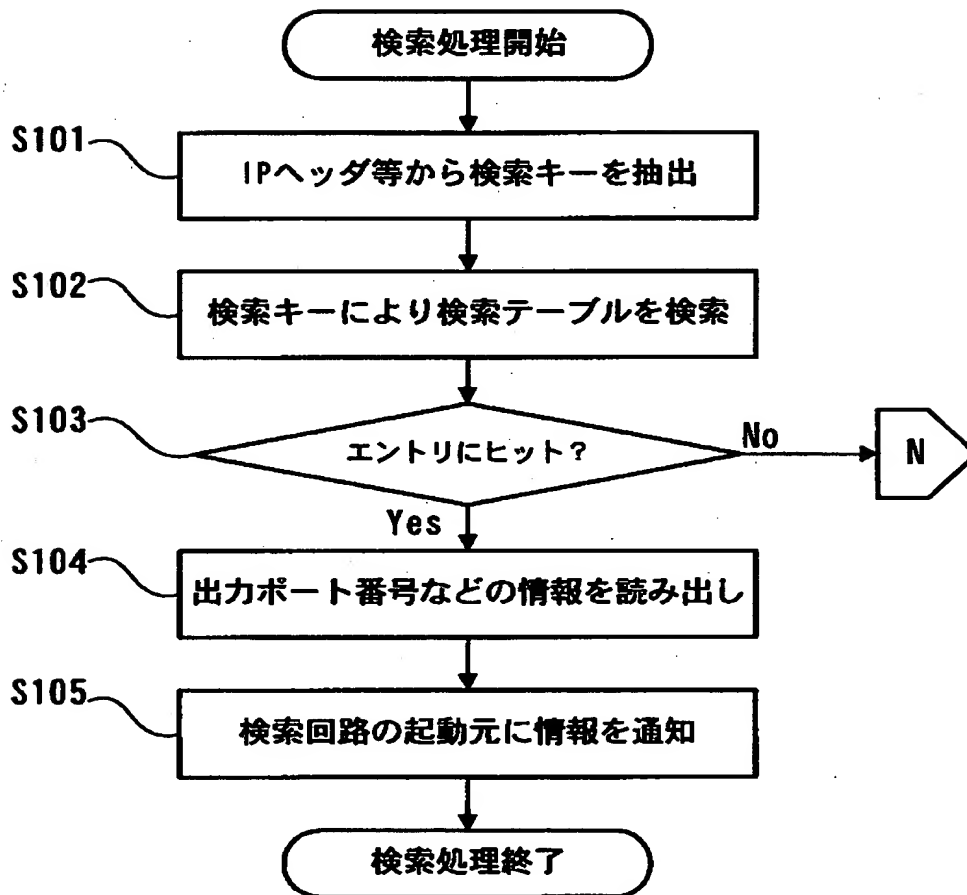
【図 2】

サーチキー			ルーティング結果			セキュリティ情報		
IP Source Address	IP Destination Address	MAC Source Address	MAC Destination Address	Output Port No.	Enryption Algorithm	Enryption Key	New IP Source Address	New IP Destination Address
AA BB CC DD	EE FF GG HH	123456	789012	2				
AA BB CC DD	KK LL MM NN			9	DES-CBC	XXXXXX	AA BB CC XX	KK LL MM YY
AA BB CC XX	KK LL MM YY	123456	789012	2				
KK LL MM YY	AA BB CC XX			9	DES-CBC	YYYYYY		
...

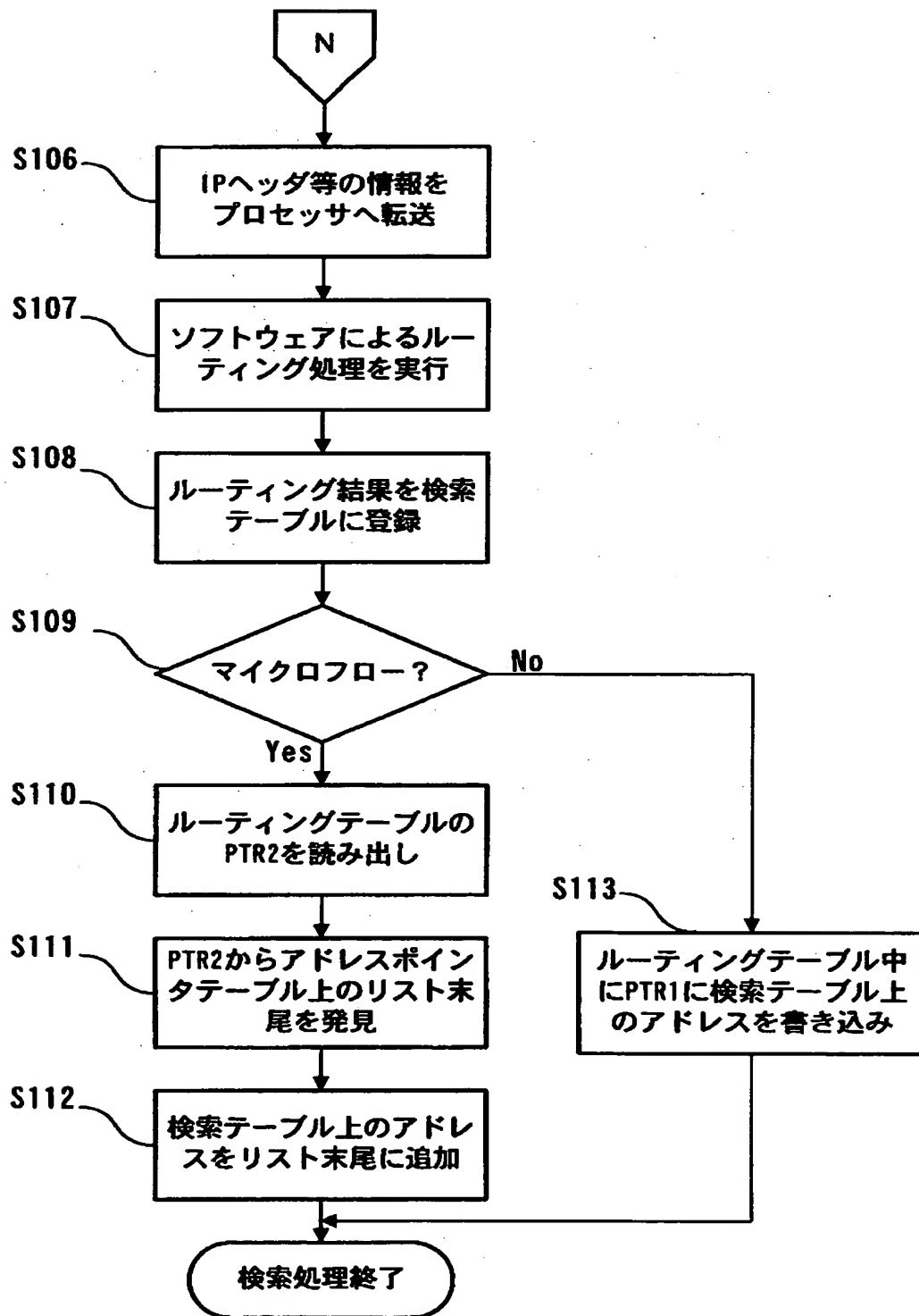
【図 3】



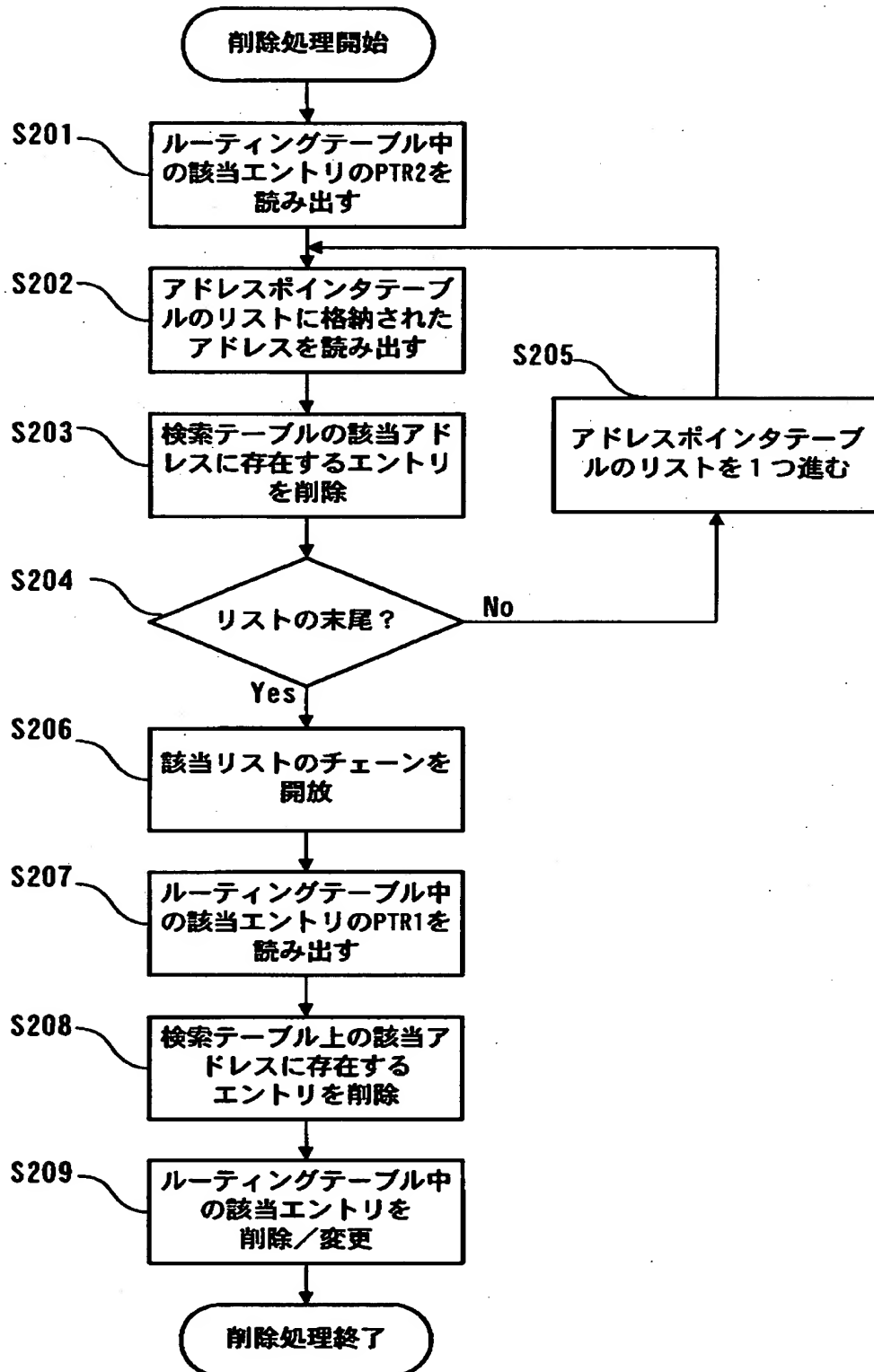
【図 4】



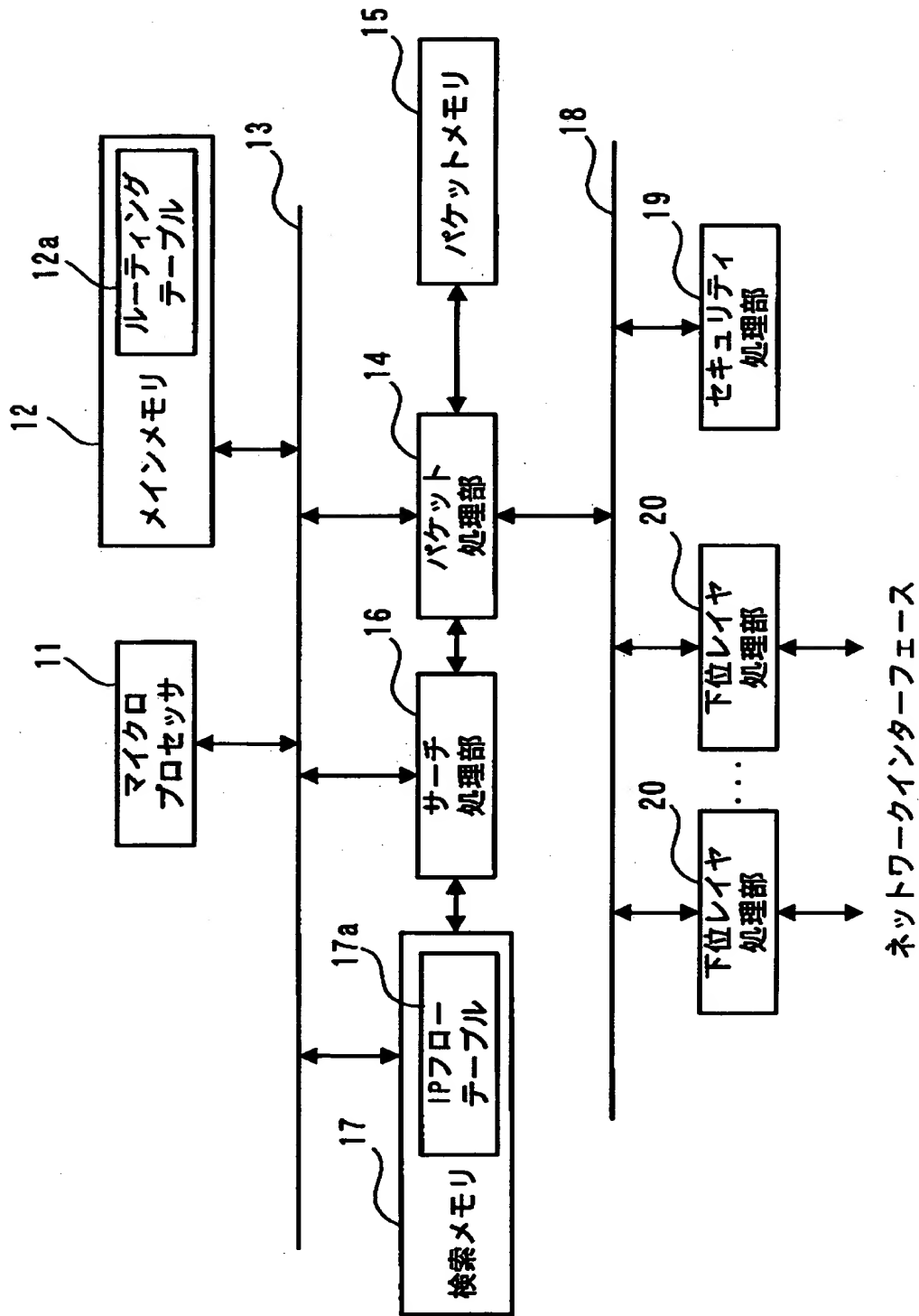
【図 5】



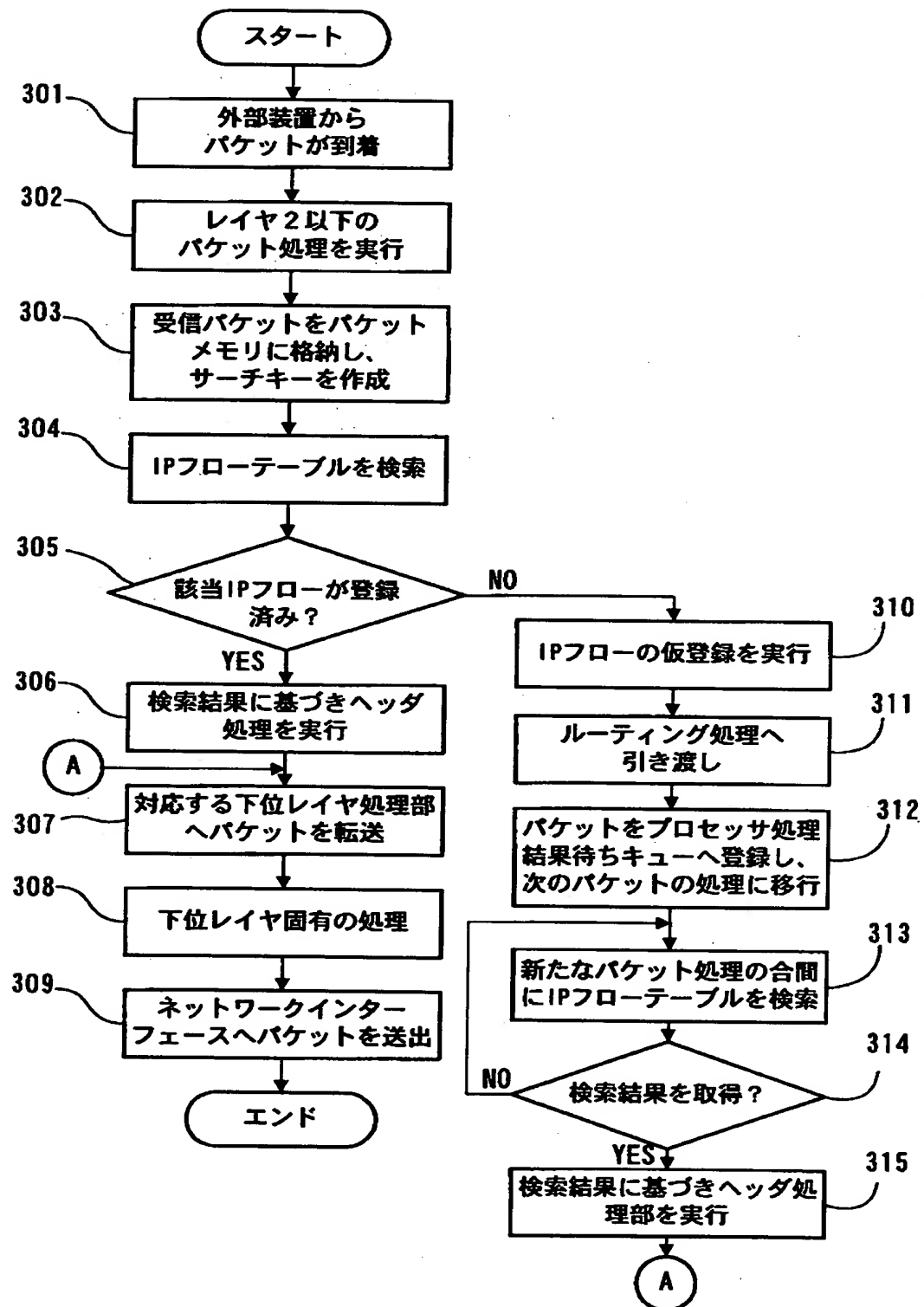
【図 6】



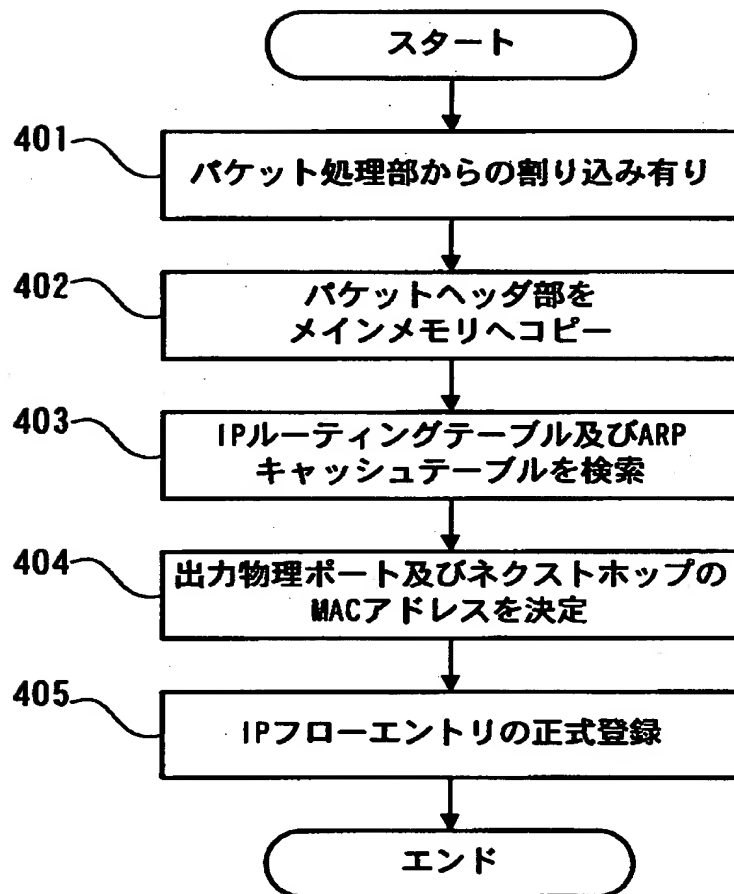
【図 7】



【図 8】



【図 9】



【書類名】 要約書

【要約】

【課題】 データベースと、このデータベースから選択・複写されたデータを記憶する検索高速化用の検索用テーブルを含み構成されたデータ検索システムの検索処理を高速化する。この種データ検索システムを含み構成されたパケット処理装置の検索用テーブルの管理処理時間を大幅に短縮し、もってパケット処理性能の低化を防止した装置と制御方法を提案する。

【解決手段】 データベースと、このデータベースから選択・複写されたデータを記憶する検索用テーブルを用いたデータ検索システムに、検索用テーブル上の各データのエントリアドレスの位置情報および／またはエントリアドレス相互の関係を記述したアドレスポインタテーブルを設ける。マイクロプロセッサと、パケット処理を高速化するための検索用テーブルが接続された検索用回路とを具備したパケット処理装置において、検索用テーブル上のエントリアドレスの位置情報および／またはエントリアドレス相互の依存関係を記述したアドレスポインタテーブルを設ける。

【選択図】 図 1

認定・付加情報

特許出願の番号	平成11年 特許願 第348422号
受付番号	59901195915
書類名	特許願
担当官	坪 政光 8844
作成日	平成11年12月17日

<認定情報・付加情報>

【提出日】 平成11年12月 8日

職権訂正履歴（職権訂正）

特許出願の番号	平成11年 特許願 第348422号
受付番号	59901195915
書類名	特許願
担当官	坪 政光 8844
作成日	平成11年12月17日

<訂正内容1>

訂正ドキュメント

明細書

訂正原因

職権による訂正

訂正メモ

【図面の簡単な説明】 および【図1】を改行した。

訂正前内容

.....。【図面の簡単な説明】 【図1】

訂正後内容

.....。

【図面の簡単な説明】

【図1】

出 願 人 履 歴 情 報

識別番号

[000004237]

1. 変更年月日

1990年 8月29日

[変更理由]

新規登録

住 所

東京都港区芝五丁目7番1号

氏 名

日本電気株式会社